AD-A114 326 TEXAS UNIV AT AUSTIN CENTER FOR CYBERNETIC STUDIES F/G 12/1
KHINCIN-KULLBACK-LEIBLER ESTIMATION WITH INEQUALITY CONSTRAINTS--ETC(U)
NOV 81 A CHARNES, W W COOPER, J TYSSEDAL N00014-75-C-0569
UNCLASSIFIED CCS-414 NL

1 OF 1
AD A
114326

END
DATE
FILMED
5-82
DTIC

CENTER FOR CYBERNETIC STUDIES

The University of Texas
Austin, Texas 78712

12

Research Report CCS 414

KHINČIN-KULLBACK-LEIBLER ESTIMATION WITH INEQUALITY CONSTRAINTS

by

A. Charnes
W.W. Cooper
J. Tyssedal*

November 1981

*University of Bergen, Norway

This research was partly supported by ONR Contract N00014-75-C-0569 with the Center for Cybernetic Studies, The University of Texas at Austin.

CENTER FOR CYBERNETIC STUDIES

A. Charnes, Director
Business-Economics Building, 203E
The University of Texas at Austin
Austin, Texas 78712
(512) 471-1821

ABSTRACT

In this paper we extend our grasp of statistical theory, duality theory, and computational convenience to the general linear inequality situation in K^2L (or minimum discrimination information) estimation by exhibiting it as the limit of a simple one parameter sequence of equality problems. Only the finite discrete distribution case is treated here.

KEY WORDS

Khinčin-Kullback-Leibler Estimation

Inferential Distribution

Minimum Discrimination Information

Inequality Constrained Distributions

INTRODUCTION

Frequently in hypothesis testing or estimation of an "inferential" distribution (in Akaike's terminology) by the Khinčin-Kullback-Leibler (or Minimum Discrimination Information) method one has information about the possible candidate distributions in the form of linear inequalities on the components of the distribution in addition to equality (moment) constraints. In MOFS 1978, Charnes, Cooper, and Seiford [1] developed a convex programming duality theory for the K^2L method with linear inequality constraints. It is especially incisive and convenient for constraints in equality (or moment) form which, further, have connections with established statistical theory as well as the analytic and computational facility of an unconstrained extremal problem in simple smooth functions.

In this paper we extend our grasp of statistical theory, duality theory, and computational convenience to the general linear inequality situation by exhibiting it as the limit of a simple one parameter sequence of equality problems. We treat only the finite discrete distribution case here, reserving the general distribution case to a forthcoming paper which also simplifies the duality theory of the "equality" case for general distributions as developed by Ben-Tal and Charnes in [2].

INEQUALITY AND APPROXIMATING EQUALITY FORMS

The dual programs for the inequality form as presented in Charnes-Cooper-Seiford [1] are

$$\max \quad v(\delta) \overset{\Delta}{=} -\delta^T \ln \left[\frac{\delta}{ec}\right]$$

(I) s.t.
$$\delta^T A^1 = b^{1T}$$
$$\delta^T A^2 + \gamma^T = b^{2T}$$
$$\delta, \gamma \geq 0$$

and

(II)
$$\inf \xi(z) \overset{\Delta}{=} c^T e^{A^1 z^1 + A^2 z^2} - b^{1T} z^1 - b^{2T} z^2$$

s.t.
$$- z^2 \geq 0$$

To obtain the dual programs for our approximating (weighted) equality form, we employ the procedure in [1]. Thus, consider

(1)
$$K(x,y,\delta,\gamma) \overset{\Delta}{=} \sum_{i \varepsilon I^1} (c_i e^{x_i} - \delta_i x_i) + \sum_{i \varepsilon I^2} (\varepsilon c_i e^{(y_i/\varepsilon)} - \gamma_i y_i)$$

where $c_i > 0$ for all i and $\varepsilon > 0$.

Minimizing with respect to x_i and y_i gives:

(2)
$$x_i^* = \ln\left(\frac{\delta_i}{c_i}\right), \ i \ \varepsilon \ I^1, \ y_i^* = \varepsilon \ln\left(\frac{\gamma_i}{c_i}\right), \ i \ \varepsilon \ I^2$$

Hence,

(3)
$$-\sum_{i \varepsilon I^1} \delta_i \ln\left(\frac{\delta_i}{c_i}\right) - \varepsilon \sum_{i \varepsilon I^2} \gamma_i \ln\left(\frac{\gamma_i}{ec_i}\right) \leq K(x,y,\delta,\gamma)$$

To decouple, as in [1], we obtain

(4)
$$\delta^T x + \gamma^T y = b^{1T} z^1 + b^{2T} z^2$$

and

(5)
$$(\delta^T, \gamma^T) \begin{bmatrix} A^1 & A^2 \\ 0 & I \end{bmatrix} \begin{bmatrix} z^1 \\ z^2 \end{bmatrix} = (b^{1T}, b^{2T}) \begin{bmatrix} z^1 \\ z^2 \end{bmatrix}$$

when we choose

(6) $x = A^1z^1 + A^2z^2 \ , \ y = z^2$

Thereby, we have the dual problems

$$\max \ v(\delta,\gamma,\varepsilon) \triangleq -\sum_{i\varepsilon I^1} \delta_i \ln\left(\frac{\delta_i}{ec_i}\right) - \varepsilon \sum_{i\varepsilon I^2} \gamma_i \ln\left(\frac{\gamma_i}{ec_i}\right)$$

(III) s.t.

$$\delta^T A^1 = b^{1T}$$

$$\delta^T A^2 + \gamma^T = b^{2T}$$

$$\delta \ , \ \gamma \geqslant 0$$

and

(IV) $\inf \ \xi(z) \triangleq c^{1T} e^{A^1z^1 + A^2z^2} - b^{1T}z^1 - b^{2T}z^2 + \varepsilon \ c^{2T} e^{(z^2/\varepsilon)}$
with z unconstrained

(Note that the c_i, $i\varepsilon I_2$, may be chosen arbitrarily.)

The duality theory of (III) and (IV) is precisely that of the equailty case in [1], as may be seen by making the change of variables to

$$\tilde{\gamma}_i \triangleq \varepsilon\gamma_i \ , \ \tilde{c}_i \triangleq \begin{bmatrix} c_i \ , \ i\varepsilon I^1 \\ \varepsilon c_i, \ i\varepsilon I^2 \end{bmatrix} .$$

Our present form is, however, more convenient for our arguments.

We now define

(7.1) $f(z) \triangleq c^{1T} e^{A^1z^1 + A^2z^2} - b^{1T}z^1 - b^{2T}z^2 \ , \ z^2 \leqslant 0$

(7.2) $g(z,\varepsilon) \triangleq c^{1T} e^{A^1z^1 + A^2z^2} - b^{1T}z^1 - b^{2T}z^2 + \varepsilon c^{2T} e^{(z^2/\varepsilon)}$

We will then have

Theorem 1:

For some $\varepsilon > 0$, $\inf_{z^2 \leq 0} f(z)$ exists if and only if $\inf_z g(z,\varepsilon)$ exists.

Further,

$$\lim_{\varepsilon \to 0} \inf_z g(z,\varepsilon) = \inf_{z^2 \leq 0} f(z)$$

Proof:

Consider $\dfrac{\partial g}{\partial z^2_i}$. It may be written as

$$(8) \qquad \frac{\partial g}{\partial z^2_i} = (c^1k)^T e^{A^1z^1 + A^2z^2} - b^2_i + z^2_i \, c_i e^{\left(z^2_i / \varepsilon\right)} \quad , \quad i \in I^2$$

where $k_r \triangleq \dfrac{\partial}{\partial z^2_i} (r^{A^2z^2})$ and $(c^1k)^T \triangleq (c^1_1k_1, \ldots, c^1_mk_m)$.

For $z^2_i > 0$ we see that by choosing ε small enough we can make $g(z,\varepsilon)$ an increasing function in that direction. Thus, in seeking a minimum we need only consider $z^2 \leq 0$. For notational simplicity in the following and "inf" with respect to z shall always also entail $z^2 \leq 0$.

For $z^2 \leq 0$ and all $\varepsilon > 0$ we have

$$(9) \qquad f(z) \leq g(z,\varepsilon) \leq f(z) + \varepsilon \sum_{i \varepsilon I^2} c_i$$

If $\inf f(z)$ exists, then, by (9), $\inf(z) \leq g(z,\varepsilon)$ so that $\inf g(z,\varepsilon)$ exists for all $\varepsilon > 0$. Further, $\inf f(z) \leq \inf g(z,\varepsilon)$. On the other hand, if $\inf g(z,\varepsilon)$ exists, then, by (9), $\inf g(z,\varepsilon) \leq f(z) + \varepsilon \sum_{i \varepsilon I^2} c_i$ so that $\inf f(z)$ exists. Further, too, $\inf g(z,\varepsilon) \leq \inf f(z) + \varepsilon \sum_{i \varepsilon I^2} c_i$.

Hence

(10.1) $\inf f(z) \leqslant \inf g(z,\varepsilon) \leqslant \inf f(z) + \varepsilon \sum_{i\varepsilon I^2} c_i$

and

(10.2) $\lim_{\varepsilon\to 0} \inf g(z,\varepsilon) = \inf f(z)$

When f(z) has a minimum at z*, again from (9) we can conclude

(11) $f(z^*) = \lim_{\varepsilon\to 0} g(z^*,\varepsilon)$ Q.E.D.

So much for the minimization side of the duality. For the K^2L side we have

Theorem 2:

The maximum in problem (I) and, as $\varepsilon \to 0$, in problem (III) is the same.

Proof:

For $0 \leqslant \gamma_i \leqslant ec_i$, we have $-c_i \leqslant \gamma_i \ln\left(\frac{\gamma_i}{ec_i}\right) \leqslant 0$, whereas for $\gamma_i > ec_i$, $\gamma_i \ln\left(\frac{\gamma_i}{ec_i}\right) > 0$. Therefore,

(12) $-\varepsilon \sum_{i\varepsilon I^2} \gamma_i \ln\left(\frac{\gamma_i}{ec_i}\right) \leqslant \varepsilon \sum_{i\varepsilon I^2} c_i$ for all $\gamma_i \geqslant 0$.

Thus,

(13) $v(\delta,\gamma,\varepsilon) \leqslant v(\delta) + \varepsilon \sum_{i\varepsilon I^2} c_i \leqslant f(z) + \varepsilon \sum_{i\varepsilon I^2} c_i$

From (13) we conclude

(14) $\sup v(\delta,\gamma,\varepsilon) \leq \sup v(\delta) + \varepsilon \sum_{i \varepsilon I^2} c_i \leq \inf f(z) + \varepsilon \sum_{i \varepsilon I^2} c_i$

But by the theory of [1],

$\sup v(\delta,\gamma,\varepsilon) = \max v(\delta,\gamma,\varepsilon) = \inf g(z,\varepsilon) \geq \inf f(z)$ by (10.1). Hence

(15) $\inf f(z) \leq \max v(\delta,\gamma,\varepsilon) \leq \sup v(\delta) + \varepsilon \sum_{i \varepsilon I^2} c_i \leq \inf f(z) + \varepsilon \sum_{i \varepsilon I_2} c_i$

Thus letting $\varepsilon \to 0$, we have

(16) $\inf f(z) \leq \lim_{\varepsilon \to 0} \max v(\delta,\gamma,\varepsilon) \leq \sup v(\delta) \leq \inf f(z)$

and

(17) $\lim_{\varepsilon \to 0} \max v(\delta,\gamma,\varepsilon) = \sup v(\delta) = \inf f(z)$

Q.E.D.

REFERENCES

1. Charnes, A., Cooper, W.W., and Seiford, L., "Extremal Principles and Optimization Dualities for Khincin-Kullback-Leibler Estimation," Mathematische Operationsforchung und Statistik, Vol. 9, No. 1, 1978, p. 21-29.

2. Charnes, A., and Ben-Tal, A., "A Dual Optimization Framework for Some Problems of Information Theory and Statistics," Journal of Problems of Control and Information Theory, Vol. 8, (5-6), p. 387-401, 1979.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: CCS 414
2. GOVT ACCESSION NO.: AD-A111 326
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Khincin-Kullback-Leibler Estimation With Inequality Constraints
5. TYPE OF REPORT & PERIOD COVERED:
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): A. Charnes, W.W. Cooper, J. Tyssedal
8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0569
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Cybernetic Studies, UT Austin, Austin, Texas 78712
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research (Code 434), Washington, D.C.
12. REPORT DATE: November 1981
13. NUMBER OF PAGES: 9
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): This document has been approved for public release and sale; its distribution is unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Khincin-Kullback-Leibler Estimation, Inferential Distribution, Minimum Discrimination Information, Inequality Constrained Distributions
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

In this paper we extend our grasp of statistical theory, duality theory, and computational convenience to the general linear inequality situation in K^2L (or minimum discrimination information) estimation by exhibiting it as the limit of a simple one parameter sequence of equality problems. Only the finite discrete distribution case is treated here.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ATE
LMED
-8